VAA-Weekly-Progress

02/04-02/11

Context

- Uncovered some results regarding centroid variation (using all keypoints)
- Progress made towards limb ratios and Dino feature clustering
- Continued work towards keypoint definitions made for future work

Goals

- Collect more side view images - Parth
  - Add in functionality to detect if the root of tail -> throat is on the same plane; Also add to allow the inverse pose of current side view images
- Continue work on centroid variance - Parth
  - Centroid - not including head and legs
- Continue work on feature extraction - Josh
  - Implementing inference model on deadcat, running on side view images, cluster to find closest species
- Continue work on limb ratio - Zian
  - Average limbs to create vectors for each species, then use cosine similarity between species
- Looking into metrics for keypoint - Shaan
- Cleaning up and Documenting All Team Code - Shaan
- Fixing biological keypoints - Claire, Ji Bing
- Redefining visible keypoints - Medha
  - Neck, root of tail, hip, shoulder

Limb Ratio Similarity

```
Top Ten Species Most Similiar to Antelope
argali sheep: 0.9954
horse: 0.9886
moose: 0.9812
dog: 0.9799
zebra: 0.9795
deer: 0.9777
sheep: 0.9776
cow: 0.9726
fox: 0.9694
```

```
Bottom Ten Species Most Similiar to Antelope
elephant: 0.8916
noisy night monkey: 0.8907
hamster: 0.8883
bobcat: 0.8763
black bear: 0.8579
alouatta: 0.8288
monkey: 0.8175
beaver: 0.8061
chimpanzee: 0.74
uakari: 0.7339
```

VIP

PURDUE UNIVERSITY

Centroid Variation Comparison

Including All keypoints:		Not including nose, 2 eyes, 4 paws, 2 shoulders, 2 knees		Not including nose, 2 eyes, 4 paws	
argali sheep	0.998723	moose	0.999674	king cheetah	0.999517
horse	0.997937	king cheetah	0.998975	argali sheep	0.998860
deer	0.994156	argali sheep	0.998952	moose	0.998709
zebra	0.992842	squirrel	0.998441	mouse	0.997490
moose	0.992537	bison	0.998173	rabbit	0.997409
giraffe	0.991767	skunk	0.997342	giraffe	0.997163
king cheetah	0.989388	zebra	0.996980	cheetah	0.997111
sheep	0.987457	mouse	0.996583	horse	0.996849
bison	0.987215	cheetah	0.996235	zebra	0.995937
fox	0.986510	deer	0.995937	bison	0.995436

VIP

PURDUE UNIVERSITY

Centroid Variation Comparison

Including All keypoints:	Not including nose, 2 eyes, 4 paws, 2 shoulders, 2 knees	Not including nose, 2 eyes, 4 paws
raccoon 0.954433 spider monkey 0.948246 monkey 0.947901 beaver 0.947896 panda 0.947079 alouatta 0.941170 chimpanzee 0.940325 gorilla 0.931223 noisy night monkey 0.928005 uakari 0.924182	monkey 0.974409 alouatta 0.972009 otter 0.971978 hamster 0.971215 chimpanzee 0.970867 spider monkey 0.967089 noisy night monkey 0.960775 gorilla 0.959019 hippo 0.957213 uakari 0.956180	hamster 0.971134 marmot 0.969968 spider monkey 0.969805 hippo 0.962562 gorilla 0.961567 alouatta 0.960921 chimpanzee 0.960412 monkey 0.954795 uakari 0.952667 noisy night monkey 0.931481

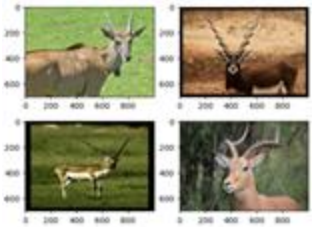
VIP

PURDUE UNIVERSITY

DINOv2 Image Feature Embeddings: Process

1. Put every side-view image(from AP10k) through the large DINOv2 Backbone

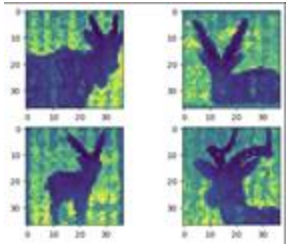
Original Images:



Quantitative Analysis:

1. Flattened the embedding of each image and then calculated the average vector for each species
2. Performed cosine similarity between the average vector of each species and the average vector of antelopes

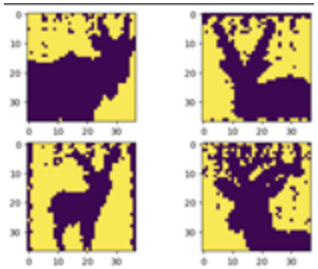
1st PCA, 1st Component:



Qualitative Analysis(only for 4 side-view Antelope images):

1. Performed PCA on the image, used the first component to isolate the background(pixels with PCA 1st component value below 0.35 are the background)
2. Performed PCA again on the foreground and made the first three components correspond to RGB(on next slide)

1st PCA, mask from 1st Component:



VIP

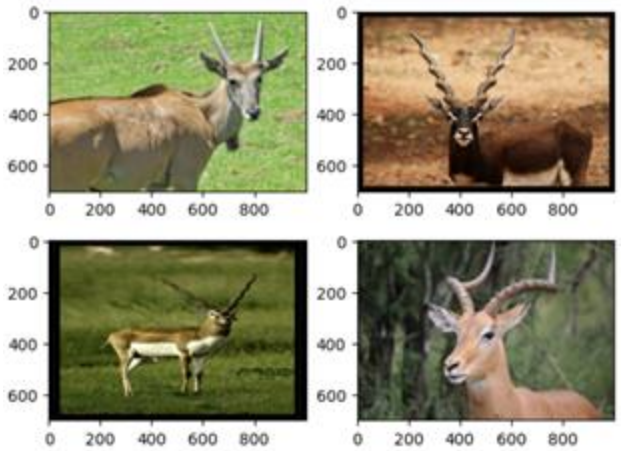
PURDUE UNIVERSITY

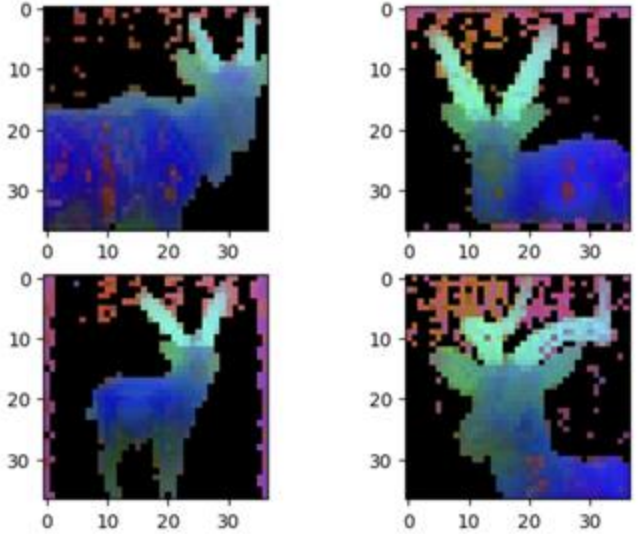
DINOv2 Image Feature Embeddings: End Results

Quantitative:

```
antelope: 1.0000
deer: 0.8048
sheep: 0.7620
moose: 0.7485
bison: 0.7443
cheetah: 0.7411
giraffe: 0.7375
fox: 0.7313
cow: 0.7251
brown bear: 0.7244
weasel: 0.7056
buffalo: 0.6940
rabbit: 0.6864
wolf: 0.6779
dog: 0.6622
raccoon: 0.6568
pig: 0.6549
skunk: 0.6499
elephant: 0.6387
squirrel: 0.6357
horse: 0.6338
hippo: 0.6278
beaver: 0.6278
argali sheep: 0.6223
mouse: 0.6188
bobcat: 0.6132
otter: 0.6100
rhino: 0.6035
snow leopard: 0.5989
polar bear: 0.5984
tiger: 0.5964
rat: 0.5862
lion: 0.5859
leopard: 0.5851
spider monkey: 0.5793
panda: 0.5764
monkey: 0.5665
jaguar: 0.5508
zebra: 0.5482
chimpanzee: 0.5346
cat: 0.5275
hamster: 0.5135
marmot: 0.5048
king cheetah: 0.4815
alouatta: 0.4501
panther: 0.4501
noisy night monkey: 0.4417
black bear: 0.4211
gorilla: 0.3279
```

Qualitative:





Visible Keypoints

- Decided to create a separate definition that prioritizes using visually-significant features (rather than being biologically correct)
- The original keypoint definition was made as a mix of biological and visible keypoints and did not have a clear objective
- In Box

Next Steps

- Finalize biological keypoint definitions and visibly distinct points, and determine other definition schemes
  - Once finalized, figure out how to divide labeling amongst group and begin labeling AP10k images
  - Determine metrics to evaluate the quality of the re-labeled AP10k keypoints
- Figure out experiment to test the similarity measures
  - Could be adding in the similar species into a training set, training RTMPose on that, and then comparing the accuracy

VIP

PURDUE UNIVERSITY

Personal Progress

Medha

- Came up with keypoint definition for visible keypoints
- Outlined the keypoint definition testing strategy

Parth

- Coded up 2 variations of centroid variance similarity measure and accumulated results
- Edited species side view extraction script to account for the inverse of left to right and making sure the tail and neck is horizontal (so no angling in the animal ideally)

Zian Pan

- Calculate the average skeleton length for each species and calculate the similarity between species by cosine similarity

Josh

- Created a script that will perform a species similarity analysis of all the species of animals in AP10k using features extracted by the the DINOv2 model, for a quantitative species similarity analysis. The script also visualizes the PCA of the first three components of the antelope, to qualitatively show that it is “learning” concepts about Antelopes.